DOCUMENT RESUME

ED 025 723 AC 003 628

By- Johnson, Raymond L.; And Others

Measuring the Educational Relevance of a Short-Term Training Program for Teachers.

Pub Date 69

Note- 17p.; Paper presented at the National Seminar on Adult Education Research (Toronto, February 9-11, 1969)

EDRS Price MF-$0.25 HC-$0.95

Descriptors- *Changing Attitudes, Housewives, *Program Evaluation, *Program Length, Questionnaires, Research, Retention, *Teacher Experience, *Teacher Workshops

A multiple time series design was used to test the effectiveness of short workshops for housewives recruited to teach basic reading to adults in bringing about changes in teacher attitudes and opinions. An attitude and opinion survey questionnaire comprising true-false items was given on three occasions, to all participants of a nine-hour workshop. All the participants answered the questionnaires at the beginning of the workshop but only the members of the experimental group answered them the second time after teaching a class. Sixteen of the 100 items showed significant shifts in responses between times one and two for the experimental group and fourteen for the control group, and virtually all the changes induced in the controls persisted at time three. However, the effects of the workshop were almost totally erased by the experiences in the classroom. (An appendix with an outline of the procedure is included). (nl)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

MEASURING THE EDUCATIONAL RELEVANCE OF A SHORT-TERM TRAINING PROGRAM FOR TEACHERS

Raymond L. Johnson
American Institutes for Research
Silver Spring, Maryland

Richard W. Cortright
Adult Education Division, NEA
Washington, D. C.

Robert S. Laubach
Laubach Literacy, Inc.
Syracuse, New York

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Raymond L. Johnson
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Presented at

the tenth National Seminar on
Adult Education Research, Toronto

February 9 - 11, 1969

MEASURING THE EDUCATIONAL RELEVANCE OF A SHORT-TERM TRAINING PROGRAM FOR TEACHERS[1]

Raymond L. Johnson, American Institutes for Research;
Richard W. Cortright, Adult Education Division, NEA;
Robert S. Laubach, Laubach Literacy, Inc.[2]

Introduction: The Practical Problems of Evaluation

When we apply to a government agency or foundation for funds to run a special training program for teachers, one of the common conditions for approval is to promise an evaluation at the end of the course to demonstrate its effectiveness. This requirement is frequently a *pro forma* one and everyone knows that some small gesture will more than satisfy the contracting officer. Very often the condition is met by appending to the final report a collection of testimonials, written by the participants on the last day of the course, all attesting to how rewarding and enriching the experience has been, and how they'll all like to come back again next year.

---

1. Presented at the National Seminar for Adult Education Research, meeting in Toronto, February 10, 1969.

2. This research was supported by a grant from the U.S. Office of Education to Laubach Literacy, Inc. An earlier study is reported by Johnson, Cortright and Cooper (1968).

The funding agencies have been so accommodating in accepting teacher testimonials as evidence of effectiveness that this approach to evaluation has almost taken on an aura of credibility. Some indication of this can be seen in a training manual published three years ago by the National Association for Public School Adult Education. The title is -- A Guide for Teacher Trainers in Adult Basic Education (1966) -- and the manual outlines in substantial detail the organization and content of a model five day training program, covering teaching methods and materials, student assessment and counseling, and the psychology of the disadvantaged adult.

Throughout the Guide, the authors stress the tentative nature of their curriculum proposals and suggestions and are quick to acknowledge that "we still know so little about how to train in this field." Yet, they also emphasize that to obtain a valid assessment of program effectiveness, data must be systematically collected over a period of time, a procedure they characterize as being both "exhaustive and exhausting." In discussing feasible alternative approaches, the Guide's section on evaluative procedures describes five methods for polling teacher trainees for their reactions to the program. Unfortunately, all the methods discussed rely entirely upon the teacher's own subjective judgments about what is relevant and useful. No method is offered for obtaining information which would permit a more impartial and

impersonal evaluation of the course.

The omission is understandable, of course, since the "classic" experimental designs for evaluation studies involve quite elaborate provisions for follow-up. An example is Withall and Fagan's (1966) evaluation of an NDEA summer institute. The investigators content analyzed tapes of classroom verbal interactions, recorded before and after the teachers attended the institute. This analysis was supplemented by extensive data obtained during in situ observations of classroom behavior. Clearly, the commitment of resources necessary for such research is prohibitive for the ordinary teacher training program. The cost of a thorough evaluation study could easily exceed the cost of the training program itself, and by a considerable margin. We're confronted here with a serious question: whether the high quality of the information gained by using a rigorous experimental design is worth its high cost, especially when the training course will have been completed long before the results of the evaluation are known.

Purpose: The Development of an Evaluative Procedure

The study we're reporting today attempted to find a compromise solution to this problem -- an approach to evaluation which will not pacify the purists in experimental design, but which nevertheless is more defensible than teacher testimonials. Our aim was to develop a relatively simple and

economical procedure of data collection and analysis suitable for the routine, moderately budgeted teacher training program.

Procedure: Sample, Survey Construction, Data Collection

Our population was a group of 63 suburban housewives who had been recruited to teach basic reading to adults. Prior to receiving their teaching assignments, they were required to attend a nine-hour training workshop, held over a three-day period, and conducted by the staff of a Milwaukee-based company which specialized in training programs. The content and organization of the course was carefully planned in advance, and we constructed a questionnaire which sampled this content. It dealt with teachers' attitudes toward adult illiterates, their opinions about student characteristics and the most successful methods of teaching and motivating them. In its format, the attitude and opinion survey consisted of 100 true-or-false items.

Each teacher trainee answered the survey on three separate occasions. Everyone completed it at the very beginning and again at the end of the nine hour workshop. We'll designate these administrations Time One and Time Two, respectively. At the conclusion of the workshop, 43 of the participants were assigned to the experimental group, and the remaining 20 were assigned to the control group. The membership of both groups was randomly determined.

Subjects in the experimental group were immediately assigned to small classes and began to teach two hours a week under close professional supervision. After ten weeks of teaching experience, the experimental group again filled out the attitude and opinion survey. This last administration we'll call Time Three.

In contrast, subjects in the control group did no teaching whatever during this period, and were assigned classes only when they had completed the survey at Time Three -- ten weeks after the workshop. They knew, of course, that they were participating in an experiment and that this was the reason for the delay.

Design: A Multiple Time Series

Giving the survey at three different times to the experimental and control groups, with teaching experience intervening between Time Two and Time Three for the experimental group, corresponds to what Campbell and Stanley (1963) have described -- in Gage's Handbook -- as a multiple time series design with a non-equivalent control group. The reason for choosing this design was that it permitted the effect of teaching experience on attitudes to be twice demonstrated: once by comparing the experimental subjects who taught with the controls who did not; and once by comparing the experimental subjects with themselves, before and after the teaching experience.

The analysis of data collected, using this design, enabled us to ask two main questions:

1. Was the short nine hour workshop effective in bringing about changes in teacher attitudes and opinions? And if so,

2. Did these changes persist, or were they merely transitory, with teachers eventually returning to their original views?

Definition of "Educational Relevance"

The answers to these two questions were important to us because they were implicated in our definition of educational relevance. The definition is based on the following premise: to be considered relevant, the content of a course should provide teachers with expectations about their future students and the teaching experience itself which are later confirmed when the teacher begins actually to deal with her students. That is to say, the changes in attitude and opinion which are induced by the course should show signs of permanence. The changes should be "locked in" by the subsequent teaching experience. This is the assumption upon which we based the evaluative procedure.

Formulating a criterion for course evaluation in terms of consistency of change observed over time enables us to set a priori performance standards for the course -- that is,

in advance, before the course is even held -- and, moreover, to state the performance standards in a way which is fully explicit.

Data Analysis: Turnover Tables, Transition Rates

The evaluative criterion requires that changes in the teachers' responses to the survey be charted over time, and this is simply done by constructing for each item a pair of turnover tables. The method is illustrated in Figure 1, on the first page of the handout (see Appendix). The turnover table on the left shows the response shifts among the 43 teachers in the experimental group, between the beginning and end of the training workshop -- that is, between Time One and Time Two. The number in the lower right hand cell corresponds to the number of teachers who believed this particular item was false at the beginning of the workshop and had not changed their minds at the end of the workshop. In the upper left hand cell is the number of teachers who, on both occasions, believed this item to be true. The two remaining cells are used to record teachers who shifted in their responses over the two observations. In the lower left hand cell are entered those who shifted from false-to-true, and in the upper right, those who shifted from true-to-false.

When we examine the marginal sums, we see that at the beginning of the workshop, the teachers' views on this item were nearly evenly split, 22 to 21. But the workshop seemed

to have some effect in increasing the number of "false" responses to 26 at Time Two.

An emergent consensus that the item was "false" is seen to have gained adherents in the turnover table on the right which shows response shifts between Time Two and Time Three. After ten weeks of classroom experience, 30 of the 43 teachers now believed the item to be false, and only 13 continued to hold the opposing view. Moreover, the strength of the minority view has steadily dwindled. Six teachers had shifted from false-to-true between Time One and Time Two, but only four had defected to the minority view between Time Two and Time Three.

It is convenient to express these shifts over time as "transition" rates, and there are a number of ways to compute them. One of the most sophisticated is a stochastic model proposed by James Coleman (1964) in his book on mathematical models in sociology. Coleman has called it a "continuous time, discrete-state model with random shocks" and it assumes that, in a state of relative equilibrium, there exist "random shocks" which are continuously moving a certain number of respondents from one cell to another in a turnover table. When the relative equilibrium is disturbed, as an effect of the experimental treatment, the result is an increase in movement -- relative to the random shock rate -- from one cell to another. The model permits exact estimates of transitions in two state turnover: that is, the transition rates

from "true-to-false" and from "false-to-true" which can be attributed to the effects of the training program and teaching experience.

The Evaluation Criterion: Two Rules Jointly Applied

Once we have calculated the two transition rates for each turnover table, we are ready to apply our evaluative criterion. The criterion consists of two rules which are jointly invoked. Rule One states that the direction of the shift observed between Time One and Time Two must remain constant from Time Two to Time Three. There must be no reversion at Time Three toward the original Time One distribution of responses. Instead, there must be evidence of a continuing shift at Time Two and Time Three to new distributions even further removed from the original one at Time One. Rule Two states that when a consistent migration is observed over time, the rate of change running counter to the dominant trend must show a corresponding reduction over time. That is, the rate of transition in a direction countervailing to the dominant trend must be inhibited.

Both of these rules find compliance in the pair of turnover tables for Figure One. The transition rates for the shift from true-to-false are larger than the shift from false-to-true, both between Time One and Time Two, and between Time Two and Time Three. This shift from true-to-false was the dominant trend and it was consistent over time. In

contrast, the rate of change running in the opposite direction shows a reduction over time. The rate of the shift from false-to-true between Time Two and Time Three is less than it was earlier, between Time One and Time Two.

For the sake of comparison, we've included an example of an item which failed to meet the joint criterion. The data are found in Figure Two on the second page of the handout. In these turnover tables, the shift from false-to-true which occurred between the beginning and end of the workshop was almost completely reversed by Time Three. The response distributions of Time One and Time Three are almost identical, and both are quite different from the distribution at Time Two. Thus, no cumulative agreement over time was observed to occur.

The two rules, jointly applied to transition rates calculated from turnover tables, make up the central elements of the evaluative procedure which is briefly outlined "cook-book fashion" in the handout.

Results: A Negative Evaluation with Positive Aspects

When the method was used to assess the effectiveness of the nine hour workshop, it became apparent that the procedure imposed a very stringent test -- that the performance standard is set quite high.

First of all, it was clear that the workshop did have a discernible if modest effect in altering teachers' attitudes and opinions. Sixteen of the 100 items showed significant

shifts in responses between Time One and Time Two for the experimental group and 14 shifted significantly in the control group. (Statistical significance here was estimated using the binomial test). Furthermore, virtually all of the changes induced in the controls were found to persist at Time Three -- a finding which suggests that even short-term courses, lasting only a few hours, are potentially useful as agents of attitude change.

We stress potentially useful because in this case the effects of the workshop were almost totally erased by the experiences in the classroom. Among teachers in the experimental group, the effect of teaching appeared to cause 14 of the 16 items which had significantly changed during the workshop to revert toward the original, pre-training distributions.

When we applied the joint evaluative criterion to the experimental group data, we found that only seven of the 100 items exhibited cumulative agreement over the three observations. And in five of the seven instances, the largest shifts occurred between Time Two and Time Three, coincident with teaching experience and not with training.

The content of the workshop must therefore be judged to have been, in large part, irrelevant -- even incompatible with what the teachers learned themselves from their contacts with adult students.

In closing, let us point to a constructive aspect of an otherwise negative evaluation of the course. The results do provide useful guidelines for revising course content and organization. Those items which showed cumulative consensus identify topics which can be elaborated upon and given greater emphasis in the future. On the other hand, those items which exhibited a strong reversal pattern help to isolate elements of course content which should be reconsidered, perhaps even omitted. So, the evaluative procedure we've described can provide some essential feedback to assist in the gradual shaping of the course into a more effective training effort.

REFERENCES

Campbell, D. T. and Stanley, J. C. "Experimental and quasi-experimental designs for research on teaching," in Handbook of Research on Teaching, N. L. Gage, ed. Chicago: Rand McNally, 1963, Chapter 5, pp. 171-246.

Coleman, J. S. Introduction to Mathematical Sociology. Glenco: Free Press, 1964, pp. 132-188.

Guide for Teacher Trainers in Adult Basic Education Washington, D. C.: National Association for Public School Adult Education, NEA, 1966.

Johnson, R. L., Cortright, R. W., and Cooper, J.V. "Attitude changes among literary teachers coincident with training and experience." Adult Education Journal, 1968, 18, 71-80.

Appendix

Measuring the Educational Relevance of A Short-Term Training Program

Outline of Procedure

1. Construct test or questionnaire to sample content of training course. Responses are dichotomous: true or false, first or second alternatives on multiple-choice type questions.

2. Administer instrument at beginning (T_1) and end (T_2) of training course, and again after the participants have acquired some post-training teaching experience (T_3).

3. Summarize responses to each item in pairs of turnover tables. Response frequencies are entered in the cells of a 2 x 2 table, with the upper left-hand and lower right-hand cells showing the number of teachers who did not change their responses to an item on the two occasions. The upper right-hand cell shows the number who said that the item was true, but subsequently changed their response to false. The lower left-hand cell shows the number who changed from false to true.

Figure 1: "Students often secretly resent their teachers' education and higher social status."

		Time 2 t	Time 2 f	
Time 1	t	11	11	22
	f	6	15	21
		17	26	

$S_{tf} = .407$

$S_{tf} = .222$

		Time 3 t	Time 3 f	
Time 2	t	9	8	17
	f	4	22	26
		13	30	

$S_{tf} = .352$

$S_{ft} = .176$

Note: $S_{tf} > S_{ft}$, at both T_1 T_2 and T_2 T_3.

S_{ft} at $T_2T_3 < S_{ft}$ at T_1 T_2

Figure 2: "Students are apt to suffer speech defects such as stuttering."

Time 1 \ Time 2	t	f	
t	23	2	25
f	12	6	18
	35	8	

$S_{tf} = .082$

$S_{ft} = .489$

Time 2 \ Time 3	t	f	
t	25	10	35
f	2	6	8
	27	16	

$S_{tf} = .440$

$S_{ft} = .089$

4. Compute "transition" rates from true-to-false (S_{tf}) and from false-to-true (S_{ft}), between T_1 T_2 and T_2 T_3. Among several alternative procedures are:

   (i) weighted proportions, discussed in R. L. Johnson, R. W. Cortright, and J. V. Cooper, "Attitude changes among literacy teachers coincident with training and experience," Adult Education Journal, 1969, 18, 71-80.

   (ii) values determined by a "continuous time, discrete-state model with random shocks," described in J. S. Coleman, Introduction to Mathematical Sociology, New York: Free Press of Glencoe, 1964, pp. 132-188.

5. The binomial test may be used to estimate whether a shift (between T_1 T_2 or T_2 T_3) was statistically significant. This step is optional.

6. Examine the pair of transition rates calculated for each item, and apply the following joint-criterion:

   (i) the direction of the shift observed between T_1 and T_2 must remain constant from T_2 to T_3. There should be no reversion at T_3 toward the original T_1 distribution of responses. Instead, there must be evidence of a continual shift at T_2 and T_3 to new distributions even further removed from the original one at T_1.

(ii) when a consistent migration was observed over time, the rate of change running counter to the dominant trend must show a corresponding reduction over time. The rate of transition in a direction countervailing to the dominant trend is inhibited.

Transition rates for turnover tables in Figure 1 show unidirectionality of change in distributions. There is a cumulative consensus on the "false" response, with a decreasing number of defections to "true."

In contrast, data in Figure 2 shows a change in direction over time, and fail to meet the joint-criterion.

An item which exhibits convergent agreement over time indicates that this information, learned in the training course, was found to be consistent with the teachers own experiences in actual instruction, and thus to have educational relevance.

ERIC Clearinghouse
FEB 13 1969
on Adult Education